Face, content, and construct validity of the *EndoViS* training system for objective assessment of psychomotor skills of laparoscopic surgeons

Fernando Pérez Escamirosa · Ricardo Manuel Ordorica Flores · Ignacio Oropesa García · Cristian Rubén Zalles Vidal · Arturo Minor Martínez

Abstract

Background The aim of this study is to present face, content, and constructs validity of the endoscopic orthogonal video system (*EndoViS*) training system and determines its efficiency as a training and objective assessment tool of the surgeons' psychomotor skills.

Methods Thirty-five surgeons and medical students participated in this study: 11 medical students, 19 residents, and 5 experts. All participants performed four basic skill tasks using conventional laparoscopic instruments and *EndoViS* training system. Subsequently, participants filled out a questionnaire regarding the design, realism, overall functionality, and its capabilities to train hand–eye coordination and depth perception, rated on a 5-point Likert scale. Motion data of the instruments were obtained by means of two webcams built into a laparoscopic physical trainer. To identify the surgical instruments in the images, colored markers were placed in each instrument. Thirteen motion-related metrics were used to assess laparoscopic performance of the participants. Statistical analysis of performance was made between novice, intermediate, and expert groups. Internal consistency of all metrics was analyzed with Cronbach's α test.

Results Overall scores about features of the *EndoViS* system were positives. Participants agreed with the usefulness of tasks and the training capacities of *EndoViS* system (score >4). Results presented significant differences in the execution of three skill tasks performed by participants. Seven metrics showed construct validity for assessment of performance with high consistency levels.

Conclusions *EndoViS* training system has been successfully validated. Results showed that *EndoViS* was able to differentiate between participants of varying laparoscopic experience. This simulator is a useful and effective tool to objectively assess laparoscopic psychomotor skills of the surgeons.

Keywords Laparoscopic surgery · Surgical training · Objective assessment · Motion metrics · Validation · Endoscopic orthogonal video system (*EndoViS*)

Laparoscopic surgery has become an important technique within several surgical specialties, such as general surgery, gynecology, and urology. This minimally invasive technique offers many benefits for the patients as less postoperative pain, better cosmetics results, and shorter periods of hospitalization [1–3]. Laparoscopic surgery, however, demands additional psychomotor abilities and skills different from those in conventional open surgery [4, 5].

Traditionally, surgical residents acquire minimally invasive skills based on the classic apprenticeship model with hands-on training in the operation room [6, 7]. This training method is not efficient, prolongs the learning curve of surgeons, and represents a potential risk to patient safety. Due to concerns for medical safety, it is essential the development of training methods for safe practice of laparoscopic surgery outside of the operating theater with the additional assessment of the surgical skills of surgeons.

Currently, laparoscopic surgical simulators are widely accepted and incorporated into surgical residency programs. These simulators have become in effective means for acquiring, training, and maintaining of psychomotor skills, which may be transferable to the operating room environment [8–11]. In the literature, there are several laparoscopic surgical simulators available for training laparoscopic skills, classified as laparoscopic box trainers [12–15], virtual reality (VR) simulators [16–19], and augmented reality (AR) simulators [20–23].

Laparoscopic box trainers are usually simple and unsophisticated simulators. These portable and inexpensive trainers allow the training of basic laparoscopic skills. However, in laparoscopic box trainers, the performance of the trainee must be evaluated by the observation of an experienced surgeon. In VR simulators, organs, tissues, and surgical laparoscopic procedures are simulated under a VR environment. These simulators provide assessment of the users without the need for an expert surgeon by means of performance-related parameters; however, most of them lack an effective and realistic haptic feedback, which decreases the realism and tactile sensation of the surgeons during training. AR simulators combine VR simulation with real images of the training modules or tasks. In contrast to VR simulators, AR simulators provide realistic haptic feedback due to the use of laparoscopic instruments, physical objects or materials, and consumables (gauze, sutures, etc.). Furthermore, these systems also offer the assessment of the efficiency of trainees using performance metrics. Nowadays, there are alternatives for the evaluation of psychomotor skills and performance using traditional laparoscopic trainers with tracking systems, which are based on mechanical, optical, or electromagnetic technologies [24]. However, their implementation might restrict the free manipulation of laparoscopic instruments altering the records and performance of the surgeon.

A viable option to these tracking systems is the video-based tracking. Video tracking systems, based on computer vision techniques, are a non-obstructive solution for capturing and analyzing of the instrument motions. In training, this approach is used in the ProMIS simulator [22, 40], in which three cameras determine the spatial position of the surgical instruments from three different angles within a mannequin. Oropesa I et al. [25] present another proposal in EVA, a tracking system that registers the 3D coordinates of the instruments based on the monoscopic image of the endoscope for assessment of skills. In general, several authors have evaluated and validated these training systems to ensure its effectiveness and usefulness using different tasks and protocols [19, 26–29].

The objective of this study is to present the *EndoViS* training system and evaluate face, content, and construct validity. *EndoViS* training system is a laparoscopic physical simulator with a video-based tracking system for evaluation of the surgical skills of surgeons during training. This training system provides a feasible alternative for surgical training programs, and its potential as a useful tool for acquisition and objective assessment of laparoscopic skills.

Materials and methods

The present study validates the *EndoViS* training system and determines the reliability of 13 motion-based metrics employed in four skill tasks. *EndoViS* system is designed and developed for learning, training, and assessment of the surgeon's psychomotor skills in laparoscopic surgery. This study was conducted in the Department of Pediatric Surgery at Hospital Infantil de México Federico Gómez in México City, México.

Participants

Experienced surgeons, residents, and medical students from the Hospital Infantil de Mexico Federico Gómez were invited to participate in this study. Participants were divided into three groups based on prior laparoscopic experience: medical students in their rotation through the department of pediatric surgery with no previous experience in laparoscopic surgery or laparoscopic simulators (Novices), residents in training ranging from postgraduate year PGY-2 to PGY-4 with less than ten laparoscopic procedures (Intermediates), and expert surgeons with experience in more than 100 laparoscopic procedures (Experts). Consequently, each participant was asked to complete a short questionnaire detailing demographic and age information.

EndoViS training system

The laparoscopic physical trainer consists of a semi-cylindrical cavity which simulates the patient's abdominal cavity (40-cm length × 33-cm width × 18-cm height), with several ports of entry that allow the insertion of standard laparoscopic instruments. The 0° optics laparoscope, provided by 750TVL resolution miniature square color camera (KT&C USA Inc., Los Angeles, CA), is in the center of the semi-cylinder. Surgical instruments are inserted through two other ports, located at each side of the optical port. Different training models may be placed on the trainer for training and assessment of skills [13, 14].

The back side of the semi-cylinder contains two webcams placed in orthogonal configuration at a distance of 13.5 cm each other that identify the laparoscopic instruments inside the simulated cavity. The orthogonal camera system captures the instruments motion in the x, y, and z planes, with a resolution of 0.14 mm and an acquisition rate of 30 frames per second (fps) [30]. The laparoscopic instruments are detected by the system using green and blue markers placed near the distal end of each surgical instrument. The images from the webcams are processed online using software developed in C language and the OpenCV library in a laptop (Asus, 2.50 GHz Intel Core i5 processor, running Windows 7 Home Basic with 6 GB RAM, 500 GB Hard Drive and 2 USB 3.0 ports). The software saves the motion data of the laparoscopic instruments in a text file (*.txt). The *EndoViS* training system was installed in the simulation lab at the hospital and placed at a standard height to perform all laparoscopic tasks (Fig. 1).

Fig. 1 *EndoViS* training system registers the laparoscopic instruments motion using an orthogonal camera system and video image processing

Face and content validity

All participants completed a questionnaire to assess face and content validity after performing the tasks with the *EndoViS* training system. The questionnaire consisted of 13 statements. The first six questions were related to the design, realism of the cavity, and functionality of the simulator (face validity), and the last seven were related to training capacities of the *EndoViS* system and the performed skill tasks (content validity). These questions were answered with a 5-point Likert scale ranging from 1 (very bad/useless) to 5 (excellent/very useful).

Construct validity

In order to evaluate construct validity of the *EndoViS* training system, all participants carried out a serial of four skill tasks.

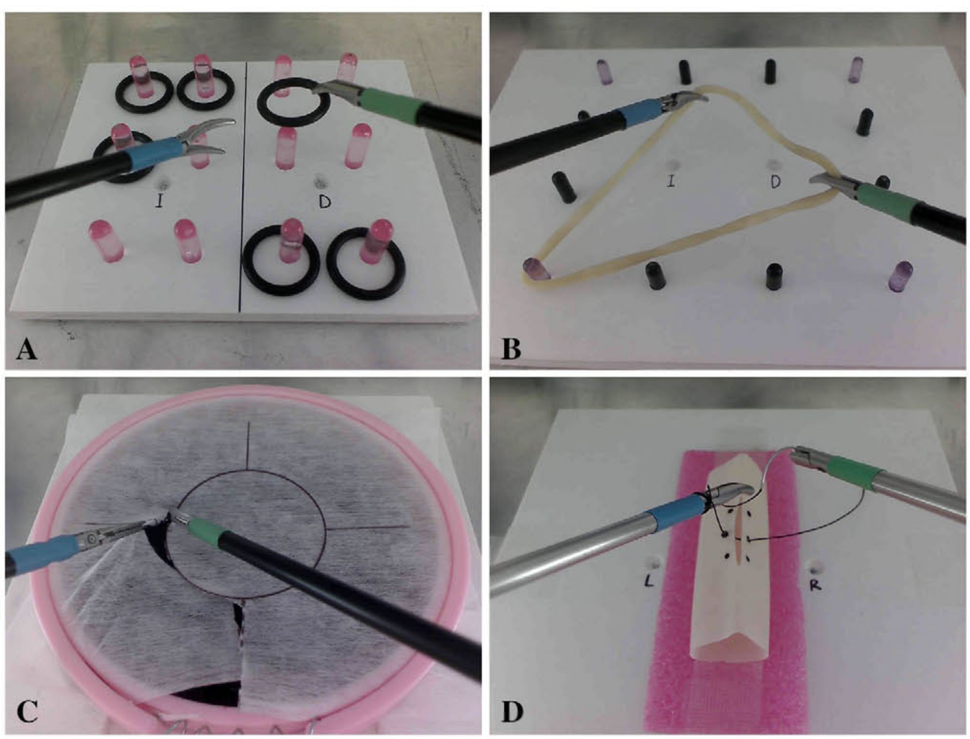

Fig. 2 Skill tasks performed by the participants. **A** Peg transfer task, **B** Rubber band task, **C** Pattern cutting task, and **D** Intracorporeal knot suture task

- *Peg transfer*: The task consisted of lifting each of six rubber rings from one peg with the dominant hand, transferring it to the non-dominant hand, and then placing it on a second peg on the opposite side of a plastic board using the laparoscopic graspers (Fig. 2A). This task involves skills at bimanual manipulation, grasping, hand–eye coordination, and spatial perception.
- *Rubber band*: The task required stretching an elastic band around 12 plastic poles placed on a plastic base (Fig. 2B). In this task, a specific order to stretch the elastic band into the posts was not defined. The participants required application of grasping, pulling force, and bimanual manipulation.
- *Pattern cutting*: The participants cut a 4.5-cm circular pattern on a piece of 13 × 13 cm nonwoven fabric stretched in a plastic base, (Fig. 2C). Using the laparoscopic scissors in his/her dominant hand, the participant cut the drawn circle as close as possible. The task ended when the circle was completely cut out and separated from the fabric. This exercise required skills at cutting, grasping, precision, and hand–eye coordination.
- *Intracorporeal knot suture*: The tasks consisted of grasping the suture needle with the laparoscopic needle driver, puncturing, and knotting a 12-cm-long suture through two predefined points in a longitudinally slit Penrose drain (Fig. 2D). The suture was tied using an intracorporeal knot technique. For all trials, 2–0 silk suture on a 26-mm taper needle was used. This task involved skills at needle manipulation, management of a silk suture, knot tying, and bimanual dexterity.

The initial and final position of the instruments was indicated by two drilled holes on the plastic boards. The placement of the tasks inside the trainer and the position of the camera were the same for all participants. Before starting the task, each participant was briefly instructed on how to perform it. All participants performed the one trial per task, and a limit of time was not imposed to complete all them.

Table 1 Summary of *EndoViS* motion-related metrics

Metrics	Definition	Equation
Time (T)	The total time required to perform the task(s) (in s)	T
Bimanual dexterity (BD)	The correlation between the velocities of the both instruments during the task(s) (–)	$\frac{\sum_{n=1}^{N}(v_{\text{left}}(n)-\bar{v}_{\text{left}})(v_{\text{right}}(n)-\bar{v}_{\text{right}})}{\sqrt{\sum_{n=1}^{N}(v_{\text{left}}(n)-\bar{v}_{\text{left}})^2\sum_{n=1}^{N}(v_{\text{right}}(n)-\bar{v}_{\text{right}})^2}}$
Path length (PL)	Total path followed by the tip of the instrument while performing the task(s) (in m)	$\int_{t=0}^{T}\sqrt{\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2+\left(\frac{dz}{dt}\right)^2}\,dt$
Depth perception (DP)	Total distance traveled by the instrument along its axis (in m)	$\int_{t=0}^{T}\sqrt{\left(\frac{dy}{dt}\right)^2+\left(\frac{dz}{dt}\right)^2}\,dt$
Depth along trocar (D_{Tr})	Total distance traveled by the instrument with respect to the trocar's coordinate (in m)	$D_{Tr}=\sqrt{(x_{Tr}-x_n)^2+(y_{Tr}-y_n)^2+(z_{Tr}-z_n)^2}$ $\int_{t=0}^{T}\frac{dD_{Tr}}{dt}\,dt$
Motion smoothness (MS)	Abrupt changes in acceleration resulting in jerky movements of the instrument (in m/s^3)	$\sqrt{\frac{T^5}{2\cdot PL^2}\int_{t=0}^{T}\left(\left(\frac{d^3x}{dt^3}\right)^2+\left(\frac{d^3y}{dt^3}\right)^2+\left(\frac{d^3z}{dt^3}\right)^2\right)^2dt}$
Average velocity (V)	Rate of change of the position of the instrument (in mm/s)	$\frac{1}{T}\int_{t=0}^{T}\sqrt{\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2+\left(\frac{dz}{dt}\right)^2}\,dt$
Average acceleration (A)	Rate of change of the velocity of the instrument (in mm/s^2)	$\frac{1}{T}\int_{t=0}^{T}\sqrt{\left(\frac{d^2x}{dt^2}\right)^2+\left(\frac{d^2y}{dt^2}\right)^2+\left(\frac{d^2z}{dt^2}\right)^2}\,dt$
Idle time (IT)	Percentage of time where the instrument was considered still (in %)	$\frac{\lvert\Im\rvert}{T}:\Im=\frac{1}{T}\int_{t=0}^{T}\sqrt{\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2+\left(\frac{dz}{dt}\right)^2}\,dt\leq 5$
Economy of area (EoA)	Relation between the maximum surface area covered by the instrument and the total path (–)	$\frac{\sqrt{[\text{Max}(x)-\text{Min}(x)]\cdot[\text{Max}(y)-\text{Min}(y)]}}{PL}$
Economy of volume (EoV)	Relation between the maximum volume covered by the instrument and the total path (–)	$\frac{\sqrt[3]{[Max(x)-Min(x)]\cdot[\text{Max}(y)-\text{Min}(y)]\cdot[\text{Max}(z)-\text{Min}(z)]}}{PL}$
Energy of area (EA)	Energy inverted by the instrument over the surface area covered (in J/cm^2)	$\frac{\sum_{t=0}^{T}\lvert x_i\rvert^2+\sum_{t=0}^{T}\lvert y_i\rvert^2}{[\text{Max}(x)-\text{Min}(x)]\cdot[\text{Max}(y)-\text{Min}(y)]}$
Energy of volume (EV)	Energy inverted by the instrument over the volume covered. (in J/cm^3)	$\frac{\sum_{t=0}^{T}\lvert x_i\rvert^2+\sum_{t=0}^{T}\lvert y_i\rvert^2+\sum_{t=0}^{T}\lvert z_i\rvert^2}{[\text{Max}(x)-\text{Min}(x)]\cdot[\text{Max}(y)-\text{Min}(y)]\cdot[\text{Max}(z)-\text{Min}(z)]}$

Table 2 Results statements: face and content validity

	Total mean	Novices		Intermediates		Experts		p^a
		Mean	SD	Mean	SD	Mean	SD	
Face validity								
Design	4.34	4.37	0.51	4.25	0.57	4.40	0.54	0.835
Realism/imitation cavity	3.69	4.00	0.75	3.68	0.70	3.40	0.54	0.327
Trocar's position	4.20	4.37	0.91	4.43	0.72	3.80	0.83	0.286
Freedom of movement of the instruments	4.22	4.50	0.53	4.56	0.51	3.60	0.54	**0.019**
User-friendliness	4.25	3.87	0.83	4.50	0.63	4.40	0.54	0.163
Overall functionality	4.39	4.37	0.74	4.62	0.50	4.20	0.44	0.278
Content validity								
Training capacities								
Hand–eye coordination	4.56	4.87	0.35	4.43	0.72	4.40	0.89	0.292
Depth perception	4.42	4.75	0.46	4.31	0.79	4.20	0.83	0.319
Basic endoscopic procedures	4.39	4.75	0.46	4.43	0.51	4.00	0.70	0.094
Tasks								
Peg transfer	4.63	4.75	0.70	4.75	0.44	4.40	0.54	0.241
Rubber band	4.42	4.50	0.89	4.56	0.48	4.20	0.83	0.224
Pattern cutting	4.29	4.00	0.71	4.68	0.51	4.20	0.83	0.366
Intracorporeal knot suture	4.27	4.25	0.83	4.56	0.51	4.00	0.70	0.205

[a] Kruskal–Wallis test for independent samples (significance $p < 0.05$)

A series of motion-related metrics employed in the study were defined from the position $[x(t),\ y(t),\ z(t)]_{t=0}^{T}$ of the instruments derived from the *EndoViS* training system. Thirteen metrics were used for the evaluation of the performance of all the participants (Table 1) [25, 31–33]. All motion metrics were computed using Matlab Release 2012a (Mathworks, Natick, MA).

Statistical analysis

Statistical analysis was performed using SPSS v20.0 for Windows (SPSS Inc., Chicago, IL, USA). Non-parametric tests were used to compare groups. The Kruskal–Wallis test was performed between the three groups, and the Mann–Whitney U test was performed for each pair of groups. A level of $p \leq 0.05$ was considered statistically significant.

In order to analyze reliability of the metrics for the different tasks, Cronbach's α test for a standardized item was performed. A value of reliability ($\alpha > 0.7$) indicated that the metric was useful and independent of the objectives of the task. On the other hand, a value of $\alpha < 0.7$ indicated a higher degree of dependency of the metric to a specific task [34].

Results

Participants

A total of 35 surgeons, residents in training, and medical students participated in this study: five expert surgeons, 19 residents, and 11 medical students (22–50 years old; 25 male and 10 female). All participants were right handed. The expert group consisted of five male surgeons; 60 % were active in general surgery department, 20 % in gynecology, and 20 % in urology. The intermediate group consisted of 19 residents; mostly (57.8 %) were active in general surgery, 21.1 % in gynecology, and 21.1 % in urology (14 male and 5 female). The novice group consisted of 11 (31.4 %) medical students in their fifth year of study from various school of medicine in Mexico.

Face and content validity

Table 2 shows the mean values of the scores for the face validity of the *EndoViS* training system. Overall, participants rated the trainer with an average score >4 on the 5-point Likert scale (mean 4.18). The lowest mean score was given to realism/imitation of the cavity. The overall scores of the experts were lower compared with the scores of the intermediates and novices. A significant difference was found for freedom of movement of the instruments between the three groups.

In general, the training capacities of the *EndoViS* system were rated with average score >4 on the 5-point Likert scale (in Table 2). Training of hand–eye coordination and depth perception received the highest mean rating, 4.56 and 4.42, respectively. All tasks were considered useful with a mean score above of 4. The overall scores of the intermediates in this category were highest compared with the novices and experts group. No significant differences were observed between groups.

Construct validity

All participants were able to complete the four skill tasks. The results from the three skill levels for four tasks are presented in Tables 3, 4, 5 and 6. The values of dominant hand and non-dominant hand are showed separately for 11 motion metrics. Parameters as time and bimanual dexterity are independent of the hand used.

For the *peg transfer* task, statistical significant differences were found between the three groups for 11 metrics: time, path length, depth perception, depth along trocar, motion smoothness, velocity, acceleration (dominant hand), economy of area, economy of volume, energy of area (dominant hand), and energy of volume (dominant hand) (Table 3). In general, statistical differences were obtained between pairs of groups: time, path length, depth perception, and economy of area for the non-dominant hand, depth along trocar for the dominant hand, and motion smoothness for both hands. The intermediates–experts (I–E) and novices–experts (N–E) groups showed significant differences in most of the metrics. The novices–intermediates (N–I) showed lower significant differences in performance compared to the other two groups.

The *rubber band* task showed no statistical differences between the three groups (Table 4). Between the pairs of groups (N–I, I–E, N–E), no statistical differences were found for this task.

For the *pattern cutting* task, statistical significant differences were found between the three groups for eight metrics: time, path length, depth perception, depth along trocar, motion smoothness, economy of area, economy of volume, and energy of area (Table 5). Statistical differences were obtained between the three pairs of groups for three metrics: time, motion smoothness, and energy of area. The results of N–I showed statistical significant differences

Table 3 Results of motion metrics for peg transfer task; mean score (SD) and p values

Metrics	Novice	Intermediates	Expert	p^a	N–I	I–E	N–E
Time (s)	119.43 (53.50)	74.80 (24.99)	52.94 (5.31)	**0.002**	**0.023**	**0.020**	**0.002**
Bimanual dexterity (–)	0.56 (0.15)	0.57 (0.14)	0.51 (0.26)	0.660	0.843	0.376	0.495
Dominant hand							
Path length (m)	1.91 (1.03)	1.24 (0.41)	0.88 (0.10)	**0.004**	0.056	**0.013**	**0.005**
Depth perception (m)	1.44 (0.73)	0.98 (0.31)	0.69 (0.06)	**0.003**	0.082	**0.006**	**0.003**
Depth along Trocar (m)	0.93 (0.40)	0.66 (0.21)	0.45 (0.04)	**0.001**	**0.041**	**0.004**	**0.001**
Motion smoothness (m/s^3)	12,739.73 (11,723.25)	4,349.25 (3,499.64)	1,976.11 (541.68)	**0.002**	**0.023**	**0.027**	**0.002**
Velocity (mm/s)	8.55 (1.89)	8.98 (1.46)	10.79 (0.84)	**0.034**	0.441	**0.017**	**0.032**
Acceleration (mm/s^2)	11.63 (2.39)	12.13 (1.96)	14.54 (1.14)	**0.017**	0.613	**0.008**	**0.015**
Idle time (%)	39.67 (12.76)	38.13 (11.32)	32.11 (6.59)	0.272	0.741	0.135	0.172
EoA (–)	0.037 (0.012)	0.044 (0.010)	0.061 (0.006)	**0.001**	0.140	**0.002**	**0.001**
EoV (–)	0.031 (0.011)	0.038 (0.009)	0.051 (0.004)	**0.001**	0.153	**0.002**	**0.001**
Energy of area (J/cm^2)	14.73 (5.85)	13.33 (4.13)	7.99 (1.42)	**0.001**	0.676	**0.001**	**0.001**
Energy of volume (J/cm^3)	880.46 (147.37)	951.88 (308.16)	591.67 (154.95)	**0.003**	0.582	**0.002**	**0.002**
Non-dominant hand							
Path length (m)	1.89 (0.69)	1.37 (0.47)	1.00 (0.33)	**0.009**	**0.041**	**0.046**	**0.008**
Depth perception (m)	1.49 (0.53)	1.07 (0.37)	0.75 (0.25)	**0.005**	**0.033**	**0.041**	**0.005**
Depth along Trocar (m)	0.96 (0.35)	0.73 (0.24)	0.53 (0.11)	**0.013**	0.061	**0.036**	**0.015**
Motion smoothness (m/s^3)	12, 612.96 (11,244.50)	4,406.94 (3,320.70)	1,947.72 (514.09)	**0.002**	**0.029**	**0.017**	**0.002**
Velocity (mm/s)	9.70 (2.27)	10.91 (2.26)	12.88 (2.27)	**0.038**	0.210	0.068	**0.015**
Acceleration (mm/s^2)	13.60 (3.42)	15.14 (3.21)	17.74 (4.14)	0.134	0.281	0.135	0.079
Idle time (%)	34.34 (12.91)	28.22 (13.12)	19.43 (6.72)	0.102	0.344	0.158	**0.025**
EoA (–)	0.035 (0.010)	0.044 (0.011)	0.055 (0.009)	**0.004**	**0.037**	**0.036**	**0.002**
EoV (–)	0.029 (0.008)	0.037 (0.009)	0.046 (0.007)	**0.005**	0.061	**0.027**	**0.002**
Energy of area (J/cm^2)	67.00 (41.08)	51.19 (17.73)	42.01 (12.94)	0.323	0.441	0.268	0.172
Energy of volume (J/cm^3)	2,239.30 (1,113.61)	2,067.39 (916.48)	1,823.75 (789.59)	0.857	0.741	0.699	0.626

[a] Kruskal–Wallis test for differences across the three groups; significant at $p < 0.05$ (bold)

[b] Mann–Whitney U test for differences between pair of groups; significant at $p < 0.05$ (bold)

Table 4 Results of motion metrics for rubber band task; mean score (SD) and p values

Metrics	Novice	Intermediates	Expert	p^a	N–I	I–E	N–E
Time (s)	104.05 (107.90)	68.34 (27.70)	55.14 (9.20)	0.702	0.657	0.364	0.817
Bimanual dexterity (–)	0.36 (0.22)	0.50 (0.25)	0.51 (0.26)	0.333	0.165	0.809	0.247
Dominant hand							
Path length (m)	1.86 (1.95)	1.30 (0.66)	1.15 (0.073)	0.887	1.000	0.904	0.355
Depth perception (m)	1.46 (1.43)	1.04 (0.50)	0.88 (0.05)	0.909	1.000	0.904	0.418
Depth along Trocar (m)	0.88 (0.79)	0.70 (0.34)	0.60 (0.04)	0.925	0.912	0.904	0.418
Motion smoothness (m/s^3)	16,761.75 (32,357.02)	4,151.54 (3,410.55)	2,329.63 (779.66)	0.654	0.657	0.304	0.817
Velocity (mm/s)	10.93 (3.11)	10.52 (3.08)	12.97 (4.03)	0.331	0.579	0.164	0.298
Acceleration (mm/s^2)	14.87 (4.43)	14.04 (4.12)	17.80 (6.01)	0.324	0.579	0.146	0.355
Idle time (%)	32.17 (12.82)	37.82 (15.43)	29.20 (10.56)	0.243	0.291	0.116	0.728
EoA (–)	0.067 (0.034)	0.071 (0.029)	0.075 (0.005)	0.777	0.956	0.397	0.817
EoV (–)	0.048 (0.025)	0.052 (0.021)	0.050 (0.002)	0.987	0.824	0.739	0.817
Energy of area (J/cm^2)	12.30 (10.89)	8.41 (3.35)	8.99 (3.57)	.922	0.868	0.672	0.908
Energy of volume (J/cm^3)	654.75 (493.83)	450.17 (248.54)	441.33 (171.39)	0.763	0.470	0.856	0.643
Non-dominant hand							
Path length (m)	1.57 (1.46)	1.03 (0.45)	0.93 (0.24)	0.925	0.781	0.952	0.643
Depth perception (m)	1.23 (1.07)	0.82 (0.33)	0.70 (0.15)	0.771	0.617	0.628	0.563
Depth along Trocar (m)	0.81 (0.80)	0.51 (0.22)	0.47 (0.08)	0.983	0.912	0.856	1.000
Motion smoothness (m/s^3)	22,501.72 (47,262.41)	4,475.39 (4,082.26)	2,488.92 (857.48)	0.846	0.697	0.586	0.908
Velocity (mm/s)	10.80 (3.35)	9.66 (2.05)	11.85 (5.89)	0.834	0.505	0.952	0.817
Acceleration (mm/s^2)	14.35 (4.13)	12.72 (3.02)	15.58 (8.57)	0.702	0.374	0.904	0.643
Idle time (%)	36.51 (12.86)	40.02 (11.50)	38.37 (17.01)	0.889	0.560	0.928	0.908
EoA (–)	0.067 (0.031)	0.070 (0.017)	0.075 (0.011)	0.718	0.912	0.304	1.000
EoV (–)	0.051 (0.024)	0.054 (0.016)	0.055 (0.008)	0.982	0.868	0.809	1.000
Energy of area (J/cm^2)	41.93 (21.86)	50.40 (24.66)	45.98 (26.59)	0.619	0.345	0.586	0.817
Energy of volume (J/cm^3)	1,632.04 (860.80)	2,094.15 (1,292.69)	1,998.79 (1,266.42)	0.772	0.505	0.904	0.563

[a] Kruskal–Wallis test for differences across the three groups; significant at $p < 0.05$ (bold)

[b] Mann–Whitney U test for differences between pair of groups; significant at $p < 0.05$ (bold)

Table 5 Results of motion metrics for pattern cutting task; mean score (SD) and p values

Metrics	Novice	Intermediates	Expert	p^a	N–I	I–E	N–E
Time (s)	339.54 (142.24)	217.66 (92.47)	78.01 (19.51)	**0.001**	**0.039**	**0.001**	**0.004**
Path length (m)	4.03 (1.64)	3.32 (1.78)	1.24 (0.44)	**0.005**	0.266	**0.004**	**0.004**
Depth perception (m)	3.19 (1.29)	2.64 (1.40)	1.02 (0.33)	**0.004**	0.266	**0.003**	**0.004**
Depth along Trocar (m)	1.59 (0.38)	1.60 (0.85)	0.70 (0.16)	**0.013**	0.589	**0.011**	**0.004**
Motion smoothness (m/s^3)	144,905.75 (143,100.17)	50,841.02 (3,8516.51)	5,248.34 (4,202.15)	**0.001**	**0.028**	**0.002**	**0.004**
Velocity (mm/s)	5.75 (1.23)	6.58 (1.31)	6.28 (0.92)	0.473	0.266	0.906	0.291
Acceleration (mm/s^2)	8.11 (1.91)	9.35 (1.79)	8.81 (1.54)	0.476	0.240	0.611	0.570
Idle time (%)	63.24 (10.40)	56.86 (10.52)	59.60 (14.03)	0.372	.153	0.667	0.570
EoA (–)	0.019 (0.004)	0.024 (0.008)	0.046 (0.007)	**0.002**	0.266	**0.001**	**0.004**
EoV (–)	0.019 (0.005)	0.022 (0.007)	0.038 (0.004)	**0.004**	0.391	**0.003**	**0.004**
Energy of area (J/cm^2)	65.21 (22.83)	48.03 (26.25)	24.23 (4.11)	**0.002**	**0.045**	**0.005**	**0.004**
Energy of volume (J/cm^3)	1,649.97 (489.23)	1,846.15 (1,548.62)	1,382.34 (545.46)	0.755	0.576	0.845	0.465

[a] Kruskal–Wallis test for differences across the three groups; significant at $p < 0.05$ (bold)

[b] Mann–Whitney U test for differences between pair of groups; significant at $p < 0.05$ (bold)

Table 6 Results of motion metrics for intracorporeal knot suture task; mean score (SD) and p values

Metrics	Novice	Intermediates	Expert	p^{a}	N–I	I–E	N–E
Time (s)	480.82 (111.16)	217.74 (81.92)	92.82 (20.68)	**0.000**	**0.000**	**0.000**	**0.001**
Bimanual dexterity (–)	0.31 (0.10)	0.46 (0.12)	0.71 (0.12)	**0.000**	**0.005**	**0.001**	**0.001**
Dominant hand							
Path length (m)	7.77 (2.92)	3.71 (1.73)	1.47 (0.35)	**0.000**	**0.000**	**0.001**	**0.001**
Depth perception (m)	6.26 (2.44)	2.96 (1.37)	1.14 (0.27)	**0.000**	**0.000**	**0.001**	**0.001**
Depth along Trocar (m)	4.07 (1.58)	1.90 (0.90)	0.73 (0.17)	**0.000**	**0.000**	**0.001**	**0.001**
Motion smoothness (m/s^3)	192,767.83 (99,360.95)	41,852.98 (27,443.09)	6,710.23 (2,640.61)	**0.000**	**0.000**	**0.000**	**0.001**
Velocity (mm/s)	8.29 (1.52)	8.69 (1.44)	8.91 (1.21)	0.707	0.647	0.775	0.315
Acceleration (mm/s^2)	11.67 (2.31)	12.32 (2.25)	12.15 (1.95)	0.829	0.535	0.824	0.791
Idle time (%)	43.03 (9.15)	39.04 (8.20)	38.12 (8.04)	0.767	0.435	0.924	0.711
EoA (–)	0.010 (0.002)	0.022 (0.008)	0.043 (0.008)	**0.000**	**0.000**	**0.001**	**0.001**
EoV (–)	0.009 (0.001)	0.019 (0.007)	0.035 (0.006)	**0.000**	**0.000**	**0.001**	**0.001**
Energy of area (J/cm^2)	64.03 (23.22)	28.22 (16.08)	15.66 (5.94)	**0.000**	**0.001**	0.061	**0.001**
Energy of volume (J/cm^3)	2,107.92 (685.23)	1,151.61 (703.78)	962.93 (437.95)	**0.005**	**0.005**	0.757	**0.004**
Non-dominant hand							
Path length (m)	7.69 (1.91)	3.55 (1.39)	1.54 (0.49)	**0.000**	**0.000**	**0.002**	**0.001**
Depth perception (m)	5.91 (1.32)	2.82 (1.09)	1.24 (0.42)	**0.000**	**0.000**	**0.001**	**0.001**
Depth along Trocar (m)	3.96 (1.09)	1.80 (0.67)	0.82 (0.26)	**0.000**	**0.000**	**0.002**	**0.001**
Motion smoothness (m/s^3)	198, 179.84 (100,571.13)	42,709.35 (28,268.50)	6,689.45 (2,618.70)	**0.000**	**0.000**	**0.000**	**0.001**
Velocity (mm/s)	8.14 (1.08)	8.48 (1.40)	9.79 (1.82)	0.125	0.726	0.120	**0.030**
Acceleration (mm/s^2)	11.61 (1.73)	12.10 (2.32)	13.68 (2.58)	0.262	0.609	0.216	0.101
Idle time (%)	41.87 (9.79)	39.05 (6.79)	32.56 (7.34)	0.101	0.435	**0.039**	0.125
EoA (–)	0.010 (0.002)	0.020 (0.006)	0.036 (0.011)	**0.000**	**0.000**	**0.003**	**0.001**
EoV (–)	0.008 (0.001)	0.018 (0.006)	0.033 (0.009)	**0.000**	**0.000**	**0.003**	**0.001**
Energy of area (J/cm^2)	136.57 (64.30)	74.39 (34.18)	58.01 (19.30)	**0.007**	**0.009**	0.266	**0.007**
Energy of volume (J/cm^3)	3,947.58 (1,537.08)	2,455.56 (1,222.20)	2,243.06 (1,025.75)	**0.039**	**0.025**	0.634	**0.030**

[a] Kruskal–Wallis test for differences across the three groups; significant at $p < 0.05$ (bold)

[b] Mann–Whitney U test for differences between pair of groups; significant at $p < 0.05$ (bold)

in only three metrics: time, motion smoothness, and energy of area. Between the I–E and N–E groups, significant differences were found in eight metrics: time, path length, depth perception, depth along trocar, motion smoothness, economy of area, economy of volume, and energy of area.

In the *intracorporeal knot suture* task, ten metrics showed statistical significant differences between the three different groups: time, bimanual dexterity, path length, depth perception, depth along trocar, motion smoothness, economy of area, economy of volume, energy of area, and energy of volume (Table 6). The results of N–I and N–E groups demonstrated statistical differences in almost all metrics. The results of I–E showed statistical differences in 9 of the 13 metrics: time, bimanual dexterity, path length, depth perception, depth along trocar, motion smoothness, idle time (only for non-dominant hand), economy of area, and economy of volume.

Analysis of Cronbach's α showed that five metrics presented high values of reliability: time, path length, depth along trocar for dominant hand, depth perception, and motion smoothness for both hands (Table 7). In particular, motion smoothness presented a strong independence toward the tasks. Other metrics such as economy of area, economy of volume for dominant hand, path length, and depth along trocar for non-dominant hand presented medium values of reliability, without reaching the pre-established threshold. On the other hand, bimanual dexterity, velocity, acceleration, energy of volume for dominant hand, idle time, energy of area, and energy of volume for non-dominant hand showed lower values of reliability, which can be considered more dependent on tasks than other metrics.

Discussion

Surgical residency programs are increasingly focused on training and objective assessment of psychomotor skills of their trainees in the field of laparoscopic surgery. Laparoscopic simulators are effective tools for improvement of

Table 7 Cronbach's alpha test results for all metrics between the different tasks

Metrics	Reliability α
Time	**0.788**
Bimanual dexterity	0.228
Dominant hand	
Path length	**0.741**
Depth perception	**0.740**
Depth along trocar	**0.719**
Motion smoothness	**0.830**
Velocity	0.376
Acceleration	0.244
Idle time	0.421
EoA	0.682
EoV	0.630
Energy of area	0.595
Energy of volume	0.233
Non-dominant hand	
Path length	0.698
Depth perception	**0.714**
Depth along trocar	0.692
Motion smoothness	**0.859**
Velocity	0.541
Acceleration	0.520
Idle time	0.075
EoA	0.482
EoV	0.475
Energy of area	0.384
Energy of volume	0.254

Bold indicates reliability at $\alpha > 0.7$ level

laparoscopic technical skills of the surgeons under a safe learning environment for the practice before performing in the operating room. However, a validation study of the simulators is always important to determine its capacities for training and objective assessment of the surgeons' performance with different levels of experience.

The purpose of this study was to validate face, content, and construct of the *EndoViS* training system and determine the internal consistency of the motion-related metrics used in four skill tasks. The results of face and content validity showed overall positive scores. The participants considered all the items of the *EndoViS* training system from good to excellent, in particular for its overall functionality. The statement of realism/imitation of the cavity obtained acceptable scores above of 3. This outcome might reside in the limited familiarity by participants to the simulators and the lack of comparison with other training systems.

For its training capacities, the *EndoViS* system was considered by all participants as a useful tool for developing of the hand–eye coordination, depth perception, and training basic endoscope procedures, rating above of 4 on the 5-point Likert scale. The skill tasks also obtained high mean scores, rated above of 4. According to the results, all participants agreed with the usefulness and degree of difficulty that presented of the tasks selected for this study. In our study, the four skill tasks were chosen for two main reasons: (1) these tasks are well validated in many clinical studies [35–41], and (2) they contain laparoscopic skills and techniques that are usually present in many laparoscopic procedures

The results of the construct validity demonstrated that there are statistically significant differences in the execution of the tasks performed by participants with different levels of experience. The *peg transfer* task showed significant differences between novices, intermediates and experts in almost all the metrics. Similar results in the performance were found when comparing the group of novices with the group of experts and when comparing the intermediates with experts. Comparing the group of the novices with intermediates showed statistically significant differences in 6 of 13 metrics. In this task, we observed that although this task is simple to perform, the *peg transfer* shows its utility to differentiate performance between surgeons with different skill levels.

In the *pattern cutting* task, the handling of the scissors with the dominant hand was analyzed. Statistical significant differences between the three different skill levels were found for eight metrics. The group of novices vs experts and the group of intermediates versus experts obtained significant differences in their performance for eight metrics. Significant differences were found only for three metrics when comparing the group of novices with intermediates. Although only the dominant hand was analyzed in the study, the cutting task required coordination of both hands. During all trials, a grasper was used in the non-dominant hand to apply traction while the scissors, controlled by the dominant hand, cut the circle at a suitable angle and precision.

The *intracorporeal knot suture* task showed significant differences between the three skill levels of experience in 10 of 13 metrics. This task, in particular, showed better results and its potential in the evaluation of the performance between the three pairs of groups than the previous tasks. The group of novices versus intermediates and the group of novices versus experts showed statistical significant differences in ten metrics, while the group of intermediates versus experts showed significant differences in nine metrics. These results might be due to laparoscopic suturing and knot tying requiring complex movements and a similar dexterity in both hands that only intermediate surgeons and experienced surgeons dominate.

The *rubber band* task did not show any statistical difference between the three skill levels of experience and the pairs of groups. We believe that a possible reason for this may be due to the fact that each participant proposed their own strategy to stretch the elastic band into the plastic poles during development of the task. However, further studies of this task are needed to explain these findings.

In general, we found that metrics such as time, path length, depth perception, and motion smoothness are good parameters for assessment of the performance with *EndoViS* training system. These metrics have already been validated in previous studies [25, 33, 34, 36]. They also showed high reliability levels with an important degree of independence to the performed tasks. Motion smoothness presented the highest internal consistency and was considered the most independent to the skill tasks of all the metrics. A possible reason for these results may reside in the motion data processing of this training system. *EndoViS* system does not apply a post-filtering stage to the data that could dampen and lose information about the jerky movements of the instrument. Other metrics as economy of area and economy of volume showed moderate consistency levels, without reaching the desired value of reliability. Nevertheless, they proved certain utility for the evaluation of spatial dominion of the workspace using the different tasks. Moreover, metrics as bimanual dexterity, velocity, acceleration, and idle time showed the lowest reliability levels, proving them to be the most dependent on the performed tasks.

In this study, three new assessment motion-based metrics were introduced: depth along trocar, energy of area, and energy of volume. Depth along trocar researches a new aspect about depth information for skills assessment of the surgeons and is defined as the Euclidean distance between the coordinate of the trocars and the coordinate of the laparoscopic instruments tip in three-dimensional space. This new metric showed promising results of the construct validity and proved its usefulness as an evaluation metric of depth perception. Analysis of its reliability demonstrated high consistency levels and a high degree of independence toward the different tasks. Energy of area and energy of volume are designed to quantify the energy inverted by the surgeon with the instrument within the working space. Both metrics showed few significant differences and low internal consistencies during the study. Although further studies are needed to find out whether these new metrics are valid for other tasks, energy of area and energy of volume were useful in tasks requiring transfer, laparoscopic suturing, and knotting.

Finally, the current study demonstrated that the *EndoViS* training system was able to differentiate among participants of varying laparoscopic experience. The validation study showed the capacities of *EndoViS* and its usefulness as a training and objective assessment tool of psychomotor skills. *EndoViS* training system, based on computer vision techniques, offers a non-obstructive solution for tracking and analyzing the motions of laparoscopic instruments without altering the performance of the surgeons. Furthermore, due to the portability offered by the system, *EndoViS* could be included in surgical training programs for continuous education of future surgeons and in the selection of the best candidates for surgical training. Improvements of *EndoViS* system, such as design, tracking of more than two laparoscopic instruments simultaneously, exploration and automatic computation of motion metrics for evaluate the performance, and more skill tasks, will be evaluated and implemented in future studies.

Conclusion

The *EndoViS* training system has been presented and successfully validated: face, content, and construct validity. The participants of this study considered the *EndoViS* system as a useful tool for training basic laparoscopic skills as hand–eye coordination and depth perception. The results of the construct validity demonstrated the capacities of the *EndoViS* to differentiate performance between novice, intermediates, and expert groups. Significant differences were found in three of the four skills tasks and in most of the evaluated motion-based metrics. *EndoViS* training system provides a non-obstructive alternative to the traditional tracking systems and a reliable method to capture and analyze the motion of surgical instruments for objective assessment of the laparoscopic skills. This simulator has a great potential for surgical training programs as an effective training tool and continuous learning of surgical skills of future laparoscopic surgeons. Further research will be conducted using new motion metrics and skill tasks.

Acknowledgments The authors want to thank all medical students, residents, and surgeons for their enthusiastic and kindly participation in all trials and staff at Hospital Infantil de Mexico for the facilities to carry out this validation study.

Disclosures Fernando Pérez, Ricardo Ordorica, Ignacio Oropesa, Cristian Zalles, and Dr. Arturo Minor have no conflicts of interest of financial ties to disclose.

References

1. Aziz O, Constantinides V, Tekkis PP, Athanasiou T, Purkayastha S, Paraskeva P, Darzi AW, Heriot AG (2006) Laparoscopic versus open surgery for rectal cancer: a meta-analysis. Ann Surg Oncol 13(3):413–424
2. Staudacher C, Vignali A (2010) Laparoscopic surgery for rectal cancer: the state of the art. World J Gastrointest Surg 2(9):275–282. doi:10.4240/wjgs.v2.i9.275

3. Klarenbeek BR, Bergamaschi R, Veenhof AA, van der Peet DL, van den Broek WT, de Lange ES, Bemelman WA, Heres P, Lacy AM, Cuesta MA (2011) Laparoscopic versus open sigmoid resection for diverticular disease: follow-up assessment of the randomized control Sigma trial. Surg Endosc 25(4):1121–1126. doi:10.1007/s00464-010-1327-0
4. Figert PL, Park AE, Witzke DB, Schwartz RW (2001) Transfer of training in acquiring laparoscopic skills. J Am Coll Surg 193(5):533–537
5. Livingston EH, Rege RV (2004) A nationwide study of conversion from laparoscopic to open cholecystectomy. Am J Surg 188(3):205–211
6. Reznick RK (1993) Teaching and testing technical skills. Am J Surg 165(3):358–361
7. Bridges M, Diamond DL (1999) The financial impact of teaching surgical residents in the operating room. Am J Surg 177(1):28–32
8. Fried GM, Feldman LS, Vassiliou MC, Fraser SA, Stanbridge D, Ghitulescu G, Andrew CG (2004) Proving the value of simulation in laparoscopic surgery. Ann Surg 240(3):518–525
9. Cosman PH, Hugh TJ, Shearer CJ, Merrett ND, Biankin AV, Cartmill JA (2007) Skills acquired on virtual reality laparoscopic simulators transfer into the operating room in a blinded, randomised, controlled trial. Stud Health Technol Inform 125:76–81
10. Samia H, Khan S, Lawrence J, Delaney CP (2013) Simulation and its role in training. Clin Colon Rectal Surg 26(1):47–55. doi:10.1055/s-0033-1333661
11. Khan MW, Lin D, Marlow N, Altree M, Babidge W, Field J, Hewett P, Maddern G (2014) Laparoscopic skills maintenance: a randomized trial of virtual reality and box trainer simulators. J Surg Educ 71(1):79–84. doi:10.1016/j.jsurg.2013.05.009
12. Laski D, Stefaniak TJ, Makarewicz W, Proczko M, Gruca Z, Sledziński Z (2011) Structuralized box-trainer laparoscopic training significantly improves performance in complex virtual reality laparoscopic tasks. Wideochir Inne Tech Malo Inwazyjne 7(1):27–32. doi:10.5114/wiitm.2011.25666
13. Martinez AM, Espinoza DL (2007) Novel laparoscopic home trainer. Surg Laparosc Endosc Percutan Tech 17(4):300–302
14. Martinez AM, Kalach AC, Espinoza DL (2008) Millimetric laparoscopic surgery training on a physical trainer using rats. Surg Endosc 22:246–249
15. Hinata N, Iwamoto H, Morizane S, Hikita K, Yao A, Muraoka K, Honda M, Isoyama T, Sejima T, Takenaka A (2013) Dry box training with three-dimensional vision for the assistant surgeon in robot-assisted urological surgery. Int J Urol 20(10):1037–1041. doi:10.1111/iju.12101
16. Verdaasdonk EG, Stassen LP, Monteny LJ, Dankelman J (2006) Validation of a new basic virtual reality simulator for training of basic endoscopic skills: the SIMENDO. Surg Endosc 20(3):511–518
17. Woodrum DT, Andreatta PB, Yellamanchilli RK, Feryus L, Gauger PG, Minter RM (2006) Construct validity of the LapSim laparoscopic surgical simulator. Am J Surg 191(1):28–32
18. Ayodeji ID, Schijven M, Jakimowicz J, Greve JW (2007) Face validation of the Simbionix LAP Mentor virtual reality training module and its applicability in the surgical curriculum. Surg Endosc 21(9):1641–1649
19. Iwata N, Fujiwara M, Kodera Y, Tanaka C, Ohashi N, Nakayama G, Koike M, Nakao A (2011) Construct validity of the LapVR virtual-reality surgical simulator. Surg Endosc 25(2):423–428. doi:10.1007/s00464-010-1184-x
20. Stylopoulos N, Cotin S, Maithel SK, Ottensmeye M, Jackson PG, Bardsley RS, Neumann PF, Rattner DW, Dawson SL (2004) Computer-enhanced laparoscopic training system (CELTS): bridging the gap. Surg Endosc 18(5):782–789
21. Soyinka AS, Schollmeyer T, Meinhold-Heerlein I, Gopalghare DV, Hasson H, Mettler L (2008) Enhancing laparoscopic performance with the LTS3E: a computerized hybrid physical reality simulator. Fertil Steril 90(5):1988–1994. doi:10.1016/j.fertnstert.2007.08.077
22. Pellen MG, Horgan LF, Barton JR, Attwood SE (2009) Construct validity of the ProMIS laparoscopic simulator. Surg Endosc 23(1):130–139. doi:10.1007/s00464-008-0066-y
23. Botden SM, Jakimowicz JJ (2009) What is going on in augmented reality simulation in laparoscopic surgery? Surg Endosc 23(8):1693–1700. doi:10.1007/s00464-008-0144-1
24. Chmarra MK, Grimbergen CA, Dankelman J (2007) Systems for tracking minimally invasive surgical instruments. Minim Invasive Ther Allied Technol 16:328–340
25. Oropesa I, Sánchez-González P, Chmarra MK, Lamata P, Fernández A, Sánchez-Margallo JA, Jansen FW, Dankelman J, Sánchez-Margallo FM, Gómez EJ (2013) EVA: laparoscopic instrument tracking based on endoscopic video analysis for psychomotor skills assessment. Surg Endosc 27(3):1029–1039. doi:10.1007/s00464-012-2513-z
26. Van Empel PJ, Commandeur JP, van Rijssen LB, Verdam MG, Huirne JA, Scheele F, Bonjer HJ, Jeroen Meijerink W (2013) Learning curve on the TrEndo laparoscopic simulator compared to an expert level. Surg Endosc 27(8):2934–2939. doi:10.1007/s00464-013-2859-x
27. Kovac E, Azhar RA, Quirouet A, Delisle J, Anidjar M (2012) Construct validity of the LapSim virtual reality laparoscopic simulator within a urology residency program. Can Urol Assoc J 6(4):253–259. doi:10.5489/cuaj.12047
28. Ritter EM, Kindelan TW, Michael C, Pimentel EA, Bowyer MW (2007) Concurrent validity of augmented reality metrics applied to the fundamentals of laparoscopic surgery (FLS). Surg Endosc 21(8):1441–1445
29. Maithel S, Sierra R, Korndorffer J, Neumann P, Dawson S, Callery M, Jones D, Scott D (2006) Construct and face validity of MIST-VR, Endotower, and CELTS: are we ready for skills assessment using simulators? Surg Endosc 20(1):104–112
30. Pérez F, Sossa H, Martínez R, Lorias D, Minor A (2013) Video-based tracking of laparoscopic instruments using an orthogonal webcams system. World Acad Sci Eng Technol Int J Med Health Pharm Biomed Eng 7(8):184–187
31. Cotin S, Stylopoulos N, Ottensmeyer MP, Neumann P, Rattner DW, Dawson SL (2002) Metrics for laparoscopic skills trainers: the weakest link! In: MICCAI 2002. LNCS, 2488, pp. 35–43. Springer, Heidelberg
32. Chmarra MK, Kolkman W, Jansen FW, Grimbergen CA, Dankelman J (2007) The influence of experience and camera holding on laparoscopic instrument movements with the TrEndo tracking system. Surg Endosc 21:2069–2075
33. Hofstad EF, Våpenstad C, Chmarra MK, Langø T, Kuhry E, Mårvik R (2013) A study of psychomotor skills in minimally invasive surgery: what differentiates expert and nonexpert performance. Surg Endosc 27(3):854–863. doi:10.1007/s00464-012-2524-9
34. Oropesa I, Chmarra MK, Sánchez-González P, Lamata P, Rodrigues SP, Enciso S, Sánchez-Margallo FM, Jansen FW, Dankelman J, Gómez EJ (2013) Relevance of motion-related assessment metrics in laparoscopic surgery. Surg Innov 20(3):299–312. doi:10.1177/1553350612459808
35. Vassiliou MC, Ghitulescu GA, Feldman LS, Stanbridge D, Leffondré K, Sigman HH, Fried GM (2006) The MISTELS program to measure technical skill in laparoscopic surgery: evidence for reliability. Surg Endosc 20(5):744–747
36. Rivard JD, Vergis AS, Unger BJ, Hardy KM, Andrew CG, Gillman LM, Park J (2014) Construct validity of individual and summary performance metrics associated with a computer-based laparoscopic simulator. Surg Endosc 28(6):1921–1928. doi:10.1007/s00464-013-3414-5

37. Arikatla VS, Sankaranarayanan G, Ahn W, Chellali A, De S, Caroline GL, Hwabejire J, DeMoya M, Schwaitzberg S, Jones DB (2013) Face and construct validation of a virtual peg transfer simulator. Surg Endosc 27(5):1721–1729. doi:10.1007/s00464-012-2664-y
38. Chellali A, Zhang L, Sankaranarayanan G, Arikatla VS, Ahn W, Derevianko A, Schwaitzberg SD, Jones DB, Demoya M, Cao CG (2014) Validation of the VBLaST peg transfer task: a first step toward an alternate training standard. Surg Endosc. doi:10.1007/s00464-014-3538-2
39. Chmarra MK, de Klein S, Winter JC, Jansen FW, Dankelman J (2011) Objective classification of residents based on their psychomotor laparoscopic skills. Surg Endosc 24(5):1031–1039. doi:10.1007/s00464-009-0721-y
40. Van Sickle KR, McClusky DA 3rd, Gallagher AG, Smith CD (2005) Construct validation of the ProMIS simulator using a novel laparoscopic suturing task. Surg Endosc 19(9):1227–1231
41. Yamaguchi S, Yoshida D, Kenmotsu H, Yasunaga T, Konishi K, Ieiri S, Nakashima H, Tanoue K, Hashizume M (2011) Objective assessment of laparoscopic suturing skills using a motion-tracking system. Surg Endosc 25(3):771–775. doi:10.1007/s00464-010-1251-3